DOCUMENT RESUME

ED 170 305 TM 007 756

AUTHOR Frisbie, David A.
TITLE Freestanding Complex Optical Scanners.
PUB DATE Mar 78
NOTE 18p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Toronto, Ontario, Canada, March, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Electronic Data Processing; Evaluation Criteria; *Guidelines; *Optical Scanners; *Performance Criteria; *Performance Specifications; Summative Evaluation; *Test Scoring Machines

ABSTRACT

A complex freestanding optical mark recognition (OMR) scanner is one which is not on-line to an external processor; it has intelligence stemming from an internal processor located within the unit or system. The advantages and disadvantages of a complex OMR can best be assessed after identifying the scanning needs and constraints of the potential user (institution). The process of selecting a scanner from several available on the market can be facilitated by examining each unit feature-by-feature. Suggested characteristics include: (1) physical characteristics; (2) performance characteristics--production rate, optical sensing capability, forms requirements, scoring functions, accuracy, and output characteristics; and (3) general characteristics--available system configurations, ease of operation, and cost. Although complex freestanding scanners are not necessarily better than simple externally-controlled scanners, one important advantage is their ability to produce hard copy (answer sheets with scores, listings, and item analyses) quickly. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED170305

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Freestanding Complex Optical Scanners

David A. Frisbie

Measurement and Research Division

Office of Instructional Resources

University of Illinois at Urbana-Champaign

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

David A. Frisbie

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Presented as part of the symposium:

Optical Mark Readers for Testing Centers

at the Annual Meeting of the

National Council on Measurement in Education

Toronto

March 1978

Printed in the U.S.A.

Freestanding Complex Optical Scanners

Introduction

The purpose of this paper is to generally describe freestanding complex optical scanners, identify the typical strengths and weaknesses of such scanners, and outline an analytical process which might be used to determine if a freestanding complex optical mark reader (OMR) would suit the needs of a particular institution. The term "freestanding" as used here will mean the scanner is not on-line to a main-frame or external miniprocessor of any type; control is solely in the hands of the operator within the system. Though complexity can refer to many aspects of a scanning system or process, here "complex" will be used to reflect the presence of intelligence in the system which gives the system a certain amount of self-sufficiency. These terms can be contrasted with "externally controlled" and "simple", descriptors which will be the basis of the next presentation in this symposium. A freestanding complex OMR can generally be characterized by the following: independence from any external intelligent processor, high scanning speed, and a variety of programming and editing controls. Examples of OMR systems which have these characteristics and which are currently used for educational applications include NCS OpScan 17 and 100, NCS Sentry 70 models, IBM 3881 and Westinghouse Learning W-2300. Many of my remarks will relate to the NCS 7010 model and the process used in selecting this model for our work at the University of Illinois.

The Measurement and Research Division of the Office of Instructional Resources at the University of Illinois at Urbana-Champaign operates an NCS model 7010 OMR system. The 7010 reader has the optional select stacker (7011)

and transport printer (7035), all of which operate with direction from a large control cabinet. A 16-bit processor with 8K of memory, a cassette drive for program loading, and a nine track 800 BPI tape unit are contained in the control cabinet. Interaction with the system is accomplished through the keyboard printer (7072), which also serves as an additional output mode. A majority of our jobs require scanning and writing the data to magnetic tape so that final data processing can be completed by an IBM 370 computer which is located in Chicago. We are not on-line; all jobs are submitted for processing at a remote terminal, located about two blocks from our scoring center. When a job requires data analysis as well as scoring, we are generally able to meet our standard of 24-hour turnaround to our customers. Simple scoring jobs can be completed while the customer waits. This brief description of our specific circumstances will be amplified further as we examine equipment features as they relate to institutional needs and constraints.

The virtues of various OMR systems cannot be evaluated thoroughly without considering the institutional needs and constraints of the potential user. System "X" may be a very good piece of equipment, but if it is too costly, too slow, or too large in size, the potential user will regard these characteristics as disadvantages. Another potential user may look on the same characteristics as advantages. As we approached the scanner selection process, we identified our primary scanning needs to be related to processing: objective examinations, course and instructor evaluation forms, research questionnaires, and internal accounting data. Certain constraints were also identified in the beginning. Physical limitations of space and weight were considered, the lack of funds to make our data processing operation self-sufficient was apparent, and the need for quick turnaround on classroom exam

processing was recognized. Once needs and limitations were sufficiently detailed, selection characteristics were outlined as shown in Appendix A. This set of specifications will be used here to help identify those characteristics which represent unique features of freestanding complex OMRs and to show how a selection decision was made on our campus.

OMR Characteristics

It is difficult to compare and contrast freestanding complex scanners and simple externally-controlled scanners as one considers selecting a system. The approach might be similar to that facing a car buyer who is trying to decide between a Cadillac and a Honda. The advantages and disadvantages associated with a Cadillac are relative to the needs and desires of the consumer. The strengths and weaknesses of the two models can more easily be judged if we examine the two products feature-by-feature. The economic individual will want to select the product which will serve his basic needs at the lowest long-run cost. So it should be in selecting a scanner. The specifications in Appendix A can serve as a focus for judging particular features of several OMR devices under consideration. A two-way grid using specifications and specific model numbers is a useful device for summarizing the relative merits of each OMR.[1]

Physical Characteristics

Space and weight requirements must be considered concurrently with some of the important performance requirements. A single externally controlled scanner takes little room and causes no structural stress problems, but it

---

[1] Datapro Research Corporation provides a written update of EDP equipment characteristics based on its vendor surveys. Volumes from May 1977 were reviewed in preparing this paper.

is also likely to be quite slow in processing a high volume of work. Several such scanners linked to the same host may require more space and carry more weight than a simple freestanding OMR which has equivalent productivity.

Generally speaking, a scanner should have its own clean power source regardless of its complexity level. Surges of power from heaters, fans, and copy machines can cause a scanner to overlook marks or momentarily lose its memory. Several scanners operating simultaneously do not require separate electrical circuits, but the user should investigate the advantages of such an arrangement.

Some OMRs are more sensitive to extreme temperature and humidity than others. Those which depend on an internal processor (minicomputer) are probably less apt to malfunction because of climatic conditions; the conditions required for the processor are sometimes more stringent than those required for a scanner. The most sensitive OMR systems are typically equipped with an automatic power-down thermostat control. When the room temperature reaches the danger level, the operator is advised that two minutes remain to shut down the equipment. Humidity control is needed to remove static conditions which could interfere severely with document transport through the scanner.

Performance Characteristics

Production rate. High speed tends to be more characteristic of complex scanners than simple ones. Scanning rate is an impure measure of production rate in some cases, however. A fast reader that must wait for a slow card punch is actually a slow reader except when cards are not needed. When output is to magnetic tape or diskette, no slow down occurs. A scanner which reads on-line may be slowed by the need to wait for other time sharers. This

problem is obviously resolved by reading off-line and transmitting at a later time, if circumstances permit. Simple, externally-controlled scanners read from 200 and up documents per hour; complex freestanding scanners read at a rate of 1,000 and up. The typical rates for the two types vary according to form size and number of sides read. (Our NCS 7010 scans 3,000 documents per hour, both sides simultaneously, with no slow down at the tape drive.)

Productivity rates must also include time required by the operator to prepare and complete a job. The number and types of headers required vary with the system. Simple scanners tend to require a few simple headers because editing and formating output is done by the host processor. Though the same is true for some complex systems, others require headers which specify the exact read locations, locations which must be selected for missing or multiple marks, and format of the output record. The potential user should consider the variety and number of jobs which must be handled at his installation in assessing the effect of set-up time.

Optical sensing capability. The ability of a scanner to select the darkest and intended mark is perhaps among the most critical factors which affect job turnaround. If stray marks, erasures, and light marks must be cleaned clerically to obtain valid data transfer, the efficiency of the system will be lowered immeasurably. Discriminatory sensitivity is less a function of system complexity and more a function of the engineering of the read mechanism. The scanner must not only be able to read marks which exist, it must also be able to detect a valid mark or distinguish between two or more marks of varying intensity. A certain amount of intelligence is required to accomplish this feat accurately. Some systems, such as those from NCS, accomplish the comparison by assigning numbers to marks according to the

amount of light which will pass through the mark. (Some scanners use the amount of light which is reflected by the mark.) When multiple marks occur, the mark with the highest numerical value (darkest) is chosen as the valid one, as long as its value exceeds the next darkest mark by two points. Computational skill in the processor is required to make the numerical comparison between multiple marks.

Related to the intensity of a mark is the form or shape of the mark. Readers which can see a single line or an "X" in a response position are generally more accurate and acceptable than those which require the entire circle or rectangle to be filled. Marks made by ballpoint pen, soft lead pencil, or hard lead pencil must sometimes be read also. A scanner which finds a large number of marking media readable is generally more useful than one which does not. The shape and type of mark which can be read are also more a function of the type of read mechanism than a function of the complexity of the scanner.

Forms requirements. The subheadings found under this characteristic (2.3) in Appendix A have little direct relationship to scanner complexity or dependence-independence of intelligence. Document sizes acceptable by a reader should be considered in view of the applications required. Density flexibility cannot be assessed independently of applications. Design complexity perhaps relates more to cost than to applications required. Forms requirements oftentimes become one of the deciding variables between two machines when all other important variables seem to balance out.

Scoring functions. Perhaps the major advantages of complex scanners over simple scanners relate to scoring ability. Most simple scanners are not scorers; they read marks which are later interpreted and scored by the host

processor. Subtest scoring and formula scoring are both typically done by relying on the intelligence of an external processor. As a result, simple scanners do not usually print a score on a test answer sheet. A simple OMR usually has no capacity to print anything. In contrast, complex scanners usually can score subtests, use one of several available scoring formulas, and print a score (or subtest and total scores) on the answer sheet. A complex OMR can usually read all item responses and compute a test score and, finally, write all of these data to a tape. The ability to program the reader and/or processor through header sheets or previously-prepared programs read in by tape gives the user added convenience and flexibility.

The obvious implication of immediate scoring is operator convenience and customer satisfaction. The importance to customers of having a number (score) printed on an answer sheet should not be overlooked. The immediate availability of a class roster of scores seems not to be an adequate replacement for a printed score. In addition, an operator can verify the reading of a scanner by checking the printed score with his own hand scoring result. Additional advantages of the print capability will be discussed in the next section of this paper.

Accuracy. The notion of accuracy relates to errors which occur in reading and the disposition of errors in the document which are detected during scanning. In the former case, an OMR should be capable of reproducing a data record on subsequent scanning of the same document. It should also treat marks of equal intensity on the same or different documents in a consistent fashion. Scanner complexity seems not to be directly related to reproducibility or consistency, but the nature of the read mechanism probably determines how reliably the machine will perform.

There is an important difference between some simple and complex scanners in terms of how error documents are treated. Many simple scanners will not physically select an error document (missing data or multiple marks in a field). Instead, all readable data are transmitted as a record with a flag in the record which indicates an error record. In other cases, the document in error is selected out, but an error record is also transmitted. The operator must make a correction on the document, if possible, and submit it for a second scanning. Many complex scanners have a select feature which physically separates error documents without transmitting a record. Some having both print capability and intelligence can print a message or code to indicate the type of error detected. The diagnostic value of such error messages is directly proportionate to the amount of intelligence of the system. Complex freestanding scanners tend to be superior to simple externally-controlled scanners with respect to both major dimensions of accuracy.

Output characteristics. By definition, externally-controlled scanners differ from freestanding scanners in at least one output characteristic; the former is on-line to an external processor and the latter is "on-line" to an internal processor. Either system is capable of creating output records on cards, magnetic tape, cassette tape or diskette, depending on the peripheral equipment to which the processor has access.

A decided advantage of the freestanding scanner system is the ability to process data immediately in one location and provide simple output for customers on the spot. Score rosters, frequency distributions, test statistics, and simple item analysis data can be prepared rapidly using utility or other stand-alone programs. The greater the memory capacity of the processor, the more complex the type of output which can be generated. (A relatively small

package of programs written by NCS is used in our facility for obtaining the type of analyses specified above.)

General Characteristics

System configurations available. There appears to be greater flexibility in designing a system around a freestanding OMR than around an externally-controlled system. This "apparent" flexibility may vanish, however, when one considers a simple scanner plus the external processor vs. the complex scanner plus the internal processor and peripheral equipment. It is possible with either system to use multiple scanners with multiple output devices or a single scanner with multiple output devices. The memory in an internal processor generally can be expanded to some definite maximum K, one which would usually be much smaller than that found in an external processor. A line printer can be included in the configuration of either system, but the maximum speed of a printer in a freestanding system is likely to be slower (600 lpm for our NCS) than that associated with an external processor (1400 lmp at our RJE). There are obvious exceptions to each of these generalizations about OMR systems, depending on the hardware available at the site in question. For example, if the external processor is some type of minicomputer rather than an IBM 370, the internal and external processors referred to here may be quite comparable in productivity.

Ease of operation. Complex freestanding scanner systems vary in their level of complexity and, therefore, generalizations about ease of operation cannot be made accurately. By comparison, however, such systems generally require more training, experience, and data processing aptitude than a simple externally-controlled system. A general administrative problem for most scoring office centers on the classification (civil service) most appropriate

for an OMR operator. Though a clerical person maybe able to operate a simple system quite handily, a more highly skilled individual is generally needed to operate a complex scanner. The fact that a vendor offers a training program for prospective system operators may be testimonial enough regarding system operation complexity. Size of operation staff and, amount and level of supervision are both additional factors which influence the level of skill required of an operator.

Cost. So many variables influence the start-up and annual costs of a system that it makes little sense to compare systems in general. Only when one has narrowed down the viable options can cost comparisons be meaningful. Factors which might be considered in making cost estimates are outlined in section 3.3 of Appendix A. Additional costs which are less direct are included in section 4.2. Other costs may surface as the potential user examines his own particular circumstances.

Summary Comments

A complex freestanding scanner is one which is not on-line to an external processor; it has intelligence stemming from an internal processor located within the unit or system. The advantages and disadvantages of a complex OMR can best be assessed after identifying the instructional needs and constraints of the environment in which it must operate. The process of selecting a scanner from several available on the market can be facilitated by examining each unit feature-by-feature.

Differences between scanners are not necessarily due to the fact that one is simple and the other complex. Certain features or characteristics are associated with complex scanners though they may not relate to the nature of the complexity at all. Temperature and humidity requirements, for example,

may affect scanning performance in both complex and simple units. The presence of static electricity in the documents will affect both types of OMRs. High speed and greater ability to discriminate between mark intensities are features associated with complex units. Ability to score and to print on a document are also associated with these units. Both types of scanners might or might not be capable of physically selecting out error documents.

One important advantage of complex scanners is their ability to turn-around hard copy immediately. That is, answer sheets with scores on them can be handed to the customer in minutes. Simple listings and item analysis can be prepared quickly without having to wait for an external host to accept the job, process, and send results through a line printer. There is no time-sharing to create waiting periods within or between job processing.

Is a complex freestanding scanner better than a simple externally-controlled scanner? Without considering the context within which the machine will function, the simple answer must be no. Is a simple externally-controlled scanner better, then? The same answer and rational apply to this question also.

APPENDIX A

Specifications to Consider in Optical Scanning Equipment

1.0 Physical Characteristics

- 1.1 Physical dimensions of each system component
- 1.2 Space requirements
- 1.3 Weight
- 1.4 Operating conditions
  - 1.41 Electrical power requirements
    - 1.41.1 Separate transformer needed?
    - 1.41.2 Voltage requirement
  - 1.42 Temperature, humidity requirements
  - 1.43 Amount of machine noise generated

2.0 Performance Characteristics

- 2.1 Scanning rate (sheets/hour)
  - 2.11 Production rate (setup, loading, unloading)
  - 2.12 Output rate (writing, punching, transmitting)
  - 2.13 Type of feed mechanism
- 2.2 Optical sensing capability
  - 2.21 Discriminatory sensitivity
    - 2.21.1 Number of positions compared/item
    - 2.21.2 Darkest mark capability
  - 2.22 Type of mark read (pencil, ballpoint, etc.)
- 2.3 Forms requirements
  - 2.31 Paper dimensions (maximum and minimum)
  - 2.32 Paper weight, composition

2.33 Ink requirements

2.34 Printing tolerance

2.35 Density flexibility (positions/line)

2.36 Designing complexity

2.36.1 Amount of training required

2.36.2 Commercial services available

2.37 Cost

2.37.1 Design and composition

2.37.2 Printing (rate change with volume?)

2.38 Ordering efficiency

2.4 Scoring functions

2.41 Number of subtests scored per pass

2.42 Formula scoring capability

2.42.1 Change formulas across subtests?

2.43 Programming requirements

2.43.1 External program required?

2.43.2 Efficiency for machine operator

2.44 Print capability

2.44.1 Score prints on document?

2.44.2 Number of scores printed

2.45 Capable of scoring two sides simultaneously?

2.46 Number of operating modes

2.46.1 Scoring only?

2.46.2 Score print and tape or card output?

2.46.3 Score and image mode?

2.47 Ability to scan different forms simultaneously

2.5 Accuracy (error rate)

2.51 Types of internal checking

2.52 Select capability?

2.52.1 Physical separation of error sheets?

2.52.2 Error sheets written to tape or punched?

2.52.3 Error sheets printed with error message

2.52.3.1 Number of message positions available per document

2.53 Error rate in production

2.53.1 Consistency for large volume

2.53.2 Reproducibility

2.6 Output characteristics

2.61 Type of data coding required (ascii, ebcdic)

2.62 Type of characters recorded

2.62.1 Alpha only?

2.62.2 Alpha or numeric?

2.62.3 Alphanumeric?

2.63 Form of output

2.63.1 Punched cards?

2.63.2 Magnetic tape

2.63.2.1 BPI requirements (800, 1600)

2.63.3 On-line capability

2.63.3.1 Reading done on- or off-line?

2.63.4 Need for special edit program to format data for subsequent processing?

3.0 General Characteristics

- 3.1 System configurations available
  - 3.11 Number of scanners/tape (punch) unit
  - 3.12 Number of tape (punch) units/scanner
  - 3.13 Convenience of expansion over time
  - 3.14 Availability of software packages for scoring and analysis
- 3.2 Ease of operation
  - 3.21 Operator skill level required
  - 3.22 Availability of training from vendor
    - 3.22.1 Amount required
    - 3.22.2 Cost of training
- 3.3 Initial cost
  - 3.31 Type of contracts available
    - 3.31.1 Lease or purchase
    - 3.31.2 Lease-purchase
  - 3.32 New equipment (scanner, tape unit, teletype, etc.)
  - 3.33 New programming or reprogramming needs
  - 3.34 Overlap time with present equipment
  - 3.35 Changes in internal and peripheral operations
  - 3.36 Forms changes
    - 3.36.1 Loss due to unused current stock
    - 3.36.2 Cost difference between new and current forms
  - 3.37 Routine maintenance cost
  - 3.38 Compatibility with existing equipment

4.0 Vendor Characteristics

- 4.1 Reputation and corporate stability

4.11 Based on annual report information

4.12 Based on remarks from current customers

4.13 Based on relationship with sales representatives

4.2 Anticipated service

4.21 Availability of routine and emergency service

4.21.1 Ease of on-site simple repairs by operator

4.22 Quality of service

4.22.1 Efficiency

4.22.2 Thoroughness

4.23 Policy regarding "down time"

4.24 Availability of consultative service for:

4.24.1 Production problems

4.24.2 Forms design questions

4.24.3 Programming problems

4.25 Availability of back-up machine in close geographic proximity